Multi-pitch Estimation using Semidefinite Programming

Jensen, Tobias Lindstrøm; Vandenberghe, Lieven

Published in:
2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

DOI (link to publication from Publisher):
10.1109/ICASSP.2017.7952946

Publication date:
2017

Document Version
Early version, also known as pre-print

MULTI-PITCH ESTIMATION USING SEMIDEFINITE PROGRAMMING

Tobias Lindstrøm Jensen[1] *and Lieven Vandenberghe*[2]

[1]Department of Electronic Systems
Aalborg University, Denmark

[2]Department of Electrical Engineering
University of California, Los Angeles

ABSTRACT

Multi-pitch estimation concerns the problem of estimating the fundamental frequencies (pitches) and amplitudes/phases of multiple superimposed harmonic signals with application in music, speech, vibration analysis, and other fields. In this paper we formulate a complex-valued multi-pitch estimator via a semidefinite programming method for continuous sparse optimization over an infinite dictionary of vectors of complex exponentials and extend this to real-valued data via a real semidefinite program with the same dimensions (*i.e.* half the size). We further impose a continuous frequency constraint naturally occurring from assuming a Nyquist sampled signal by adding an additional semidefinite constraint. In our numerical experiments, the proposed estimator shows superior performance compared to state-of-the-art methods for separating two closely spaced fundamentals and approximately achieves the asymptotic Cramér-Rao lower bound.

Index Terms— Spectral estimation, multi-pitch estimation, continuous sparse optimization, semidefinite programming

1. INTRODUCTION

Multi-pitch estimation concerns the problem of estimating the fundamental frequencies (pitches) and amplitudes/phases of multiple superimposed harmonic sources for general signal operations such as source separation, enhancement, modification and classification [1]. Multi-pitch estimation is in general complicated by the spectral smearing caused by closely spaced harmonics [2].

The multi-pitch estimation problem is for obvious reasons closely linked to line spectral estimation. A well-founded estimator for both the multi-pitch and line spectral model is the non-linear least-squares (NLS) estimator but the NLS objective function is multi-modal in both cases which limits the practical usage of the NLS estimator, see e.g. [3, 4]. To this end many approximate methods have been proposed. Several multi-pitch estimators can be seen as extensions to classical statistical signal processing methods for line spectral estimation such as the subspace method MUSIC, Capon filtering [4] and the EM algorithm [2] for multi-pitch estimation.

The work of T.L. Jensen was supported by the Danish council for Independent Research | Technology and Production grant no. 4005-00122.

Often, a complex-valued model is considered due to the simplicity of the model, and often computationally simpler algorithms, even though a real-valued model in many applications would be more appropriate [4]. Though the multi-pitch estimation problem presents itself as a continuous problem, a possible convex optimization approach is based on a discretized frequency grid with well chosen penalty functions to ensure the harmonic structure [5].

A recent framework has been studied in superresolution, gridless compressed sensing, and atomic norm optimization where it is possible to work with a continuous dictionary of complex exponentials via semidefinite programming (SDP), see e.g. [6–9]. It is interesting to note that such a superresolution formulation for line spectral estimation can also be interpreted as an approximation of the NLS estimator [3]. Various extensions of these formulations were presented in [10, 11]. Applications include line spectral estimation [9], direction-of-arrival estimation [12] and compressed sensing [8].

In this paper we will formulate a real-valued multi-pitch estimator with a frequency constraint by exploiting the model flexibility in the superresolution framework. Interestingly, the real-valued model is in this case actually as simple as the complex-valued model. Monte Carlo simulations show that the proposed estimator outperforms state-of-the-art estimators for closely spaced fundamental frequencies and approximately achieves the Cramér-Rao lower bound (CRLB).

Notation: the set of complex numbers is $\mathbb{C}$ and real numbers $\mathbb{R}$. $\mathbb{H}^n \subseteq \mathbb{C}^{n\times n}$ is the set of Hermitian matrices, $\mathbb{S} \subseteq \mathbb{R}^{n\times n}$ is the set of real symmetric matrices and $\mathbb{T}^n \subseteq \mathbb{H}^n$ is the set of Hermitian Toeplitz matrices. For $x \in \mathbb{C}^n$, $\Re(x)$ denotes the real part and $[x]_k$ denotes the kth element. $\mathbf{tr}\, X$ denotes the trace of $X \in \mathbb{C}^{n\times n}$ and $x = \mathbf{vec}(X) \in \mathbb{C}^{mn\times 1}$ denotes the column-wise vector stacked version of $X \in \mathbb{C}^{m\times n}$. The identity matrix is $I_n \in \mathbb{R}^{n\times n}$.

2. SPARSE OPTIMIZATION

Consider an optimization problem on the form

$$\begin{array}{ll} \underset{r,a_k,c_k}{\text{minimize}} & f(\sum_{k=1}^{r} a_k c_k^H) + \sum_{k=1}^{r} \|c_k\|_2 \\ \text{subject to} & a_k \in \mathbb{A}_n,\ k = 1,\ldots,r \end{array} \tag{1}$$

where f is a convex function possibly parameterized with known data and $\mathbb{A}_n \subseteq \mathbb{C}^{n\times 1}$ is a given set (or dictionary).

The unknowns in the problem are the coefficients $c_k \in \mathbb{C}^{m\times 1}, k = 1,\dots,r$, the elements/atoms $a_k, k = 1,\dots,r$ selected from $\mathbb{A}_n$ and the number of selected atoms r. When $\mathbb{A}_n$ is a finite set and $m = 1$, problem (1) includes the special case of ℓ_1-minimization (for example, LASSO or basis pursuit). Examples involving infinite sets includes nuclear norm minimization and many other applications; see [6]. In this paper we will work with a infinite/continuous dictionary $\mathbb{A}_n$ containing vectors of complex exponentials

$$\mathbb{A}_n = \left\{ \frac{1}{\sqrt{n}} z_n(\omega) \,\middle|\, |\omega - \alpha| \le \beta \right\} \tag{2}$$

$$z_n(\omega) = \left[1, \exp(j\omega), \dots, \exp(j(n-1)\omega)\right]^T . \tag{3}$$

With $\alpha = 0$, $\beta = \pi$, the problem (1) is equivalent to the SDP

$$\begin{array}{ll} \underset{X_{11},X_{12},X_{22}}{\text{minimize}} & f(X_{12}) + \frac{1}{2}(\mathbf{tr}\, X_{11} + \mathbf{tr}\, X_{22}) \\ \text{subject to} & \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^H & X_{22} \end{bmatrix} \succeq 0 \\ & X_{11} \in \mathbb{T}^n, X_{12} \in \mathbb{C}^{n\times m}, X_{22} \in \mathbb{H}^m . \end{array} \tag{4}$$

This relation was shown for $m = 1$ and various f in [7–9], and for the multiple measurement vector case $m > 1$ in [13]. From a solution $(X_{11}^\star, X_{12}^\star, X_{22}^\star)$ of (4), we have $r = \mathbf{rank}(X_{11}^\star)$ and the frequencies of the atoms $a_k = (1/\sqrt{n}) z_n(\omega_k)$ may be extracted by various methods, see e.g. [9, 11, 14].

3. THE MULTI-PITCH MODEL AND ESTIMATION

The multi-pitch model can be formulated as

$$x = \sum_{l=1}^{L} Z_K(l\omega)\bar{c}_l, \quad y = x + w \tag{5}$$

with

$$y = \begin{bmatrix} y_1 & \cdots & y_N \end{bmatrix}^T, \quad \bar{c}_l = \begin{bmatrix} \bar{c}_{l,1} & \cdots & \bar{c}_{l,K} \end{bmatrix}^T \tag{6}$$

$$\omega = \begin{bmatrix} \omega_1 & \cdots & \omega_K \end{bmatrix}^T \tag{7}$$

$$Z_K(\omega) = \begin{bmatrix} z_N(\omega_1) & \cdots & z_N(\omega_K) \end{bmatrix} \tag{8}$$

$$w = \begin{bmatrix} w_0 & \cdots & w_{N-1} \end{bmatrix}^T \sim \mathcal{CN}(0, \sigma^2 I) \tag{9}$$

where $\bar{c}_l \in \mathbb{C}^K$ are the complex amplitudes for the lth harmonic for all K sources and ω_k is the kth fundamental frequency (sometime denoted the pitch). Some elements of $\bar{c}_l$ may be zero to include varying number of harmonics for the different pitches and L should then be interpreted as the maximum harmonic model order, $L = \max\{L_1, \dots, L_K\}$.

A multi-pitch estimator may be based on the model (5) to take the data y and estimate the frequencies $\hat{\omega}_1, \dots, \hat{\omega}_K$ (the complex amplitudes $\bar{c}_l$ can be estimated from $\hat{\omega}_1, \dots, \hat{\omega}_K$ as point estimates using linear least squares). In the following, we will show how to formulate a multi-pitch estimator via a continuous sparse optimization formulation (1) with the dictionary (2) and the equivalent SDP (4).

Let $P_l \in \mathbb{R}^{N\times NL}$ be a selection matrix such that $P_l v \in \mathbb{R}^N$ is every lth element of v, $P_l v = \left[v_1, v_{1+l}, \dots, v_{1+(N-1)l}\right]$. Then

$$z_N(l\omega_k) = P_l z_{NL}(\omega_k) . \tag{10}$$

We also form the selection and add matrix

$$P = \begin{bmatrix} P_1 & P_2 & \cdots & P_L \end{bmatrix} \in \mathbb{R}^{N\times NL^2} . \tag{11}$$

Further, let $c_k = \sqrt{n}\begin{bmatrix} [\bar{c}_1]_k & \cdots & [\bar{c}_L]_k \end{bmatrix}^H \in \mathbb{C}^{L\times 1}$. With these definitions we can describe the multi-pitch signal x in (5) as

$$\sum_{l=1}^{L} Z_K(l\omega)\bar{c}_l = \sum_{l=1}^{L}\sum_{k=1}^{K} z_N(l\omega_k)[\bar{c}_l]_k \tag{12}$$

$$= \sum_{k=1}^{K}\sum_{l=1}^{L} P_l z_{NL}(\omega_k)[\bar{c}_l]_k \tag{13}$$

$$= P\,\mathbf{vec}\left(\sum_{k=1}^{K} (1/\sqrt{n})\, z_{NL}(\omega_k) c_k^H\right) \tag{14}$$

$$= P\,\mathbf{vec}\left(\sum_{k=1}^{r} a_k c_k^H\right) \tag{15}$$

for the atoms $a_k = (1/\sqrt{n})\, z_{NL}(\omega_k) \in \mathbb{A}_{NL}$ pertaining to the continuous dictionary (2) and $r = K$. A complex-valued multi-pitch estimator can then be formulated via the continuous sparse optimization problem

$$\begin{array}{ll} \underset{K,\omega_k,c_k}{\text{minimize}} & \sum_{k=1}^{K} \|c_k\|_2 \\ \text{subject to} & \|y - x\|_2 \le \delta,\ x = \sum_{l=1}^{L}\sum_{k=1}^{K} z_N(l\omega_k)[c_k]_l \\ & |\omega_k| \le \pi, k = 1,\dots,K \end{array} \tag{16}$$

with the equivalent SDP

$$\begin{array}{ll} \underset{X_{11},X_{12},X_{22}}{\text{minimize}} & \frac{1}{2}(\mathbf{tr}(X_{11}) + \mathbf{tr}(X_{22})) \\ \text{subject to} & \|y - x\|_2 \le \delta,\ x = P\,\mathbf{vec}(X_{12}) \\ & \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^H & X_{22} \end{bmatrix} \succeq 0 \\ & X_{11} \in \mathbb{T}^{NL}, X_{12} \in \mathbb{C}^{NL\times L}, X_{22} \in \mathbb{H}^L. \end{array} \tag{17}$$

A model order estimate $\hat{K}$ may be extracted from the solution as $\hat{K} = \mathbf{rank}(X_{11}^\star)$. However, in this formulation L is included and it is necessary to know (or estimate) an appropriate regularization parameter δ.

4. THE REAL-VALUED MULTI-PITCH ESTIMATOR

The real-valued model is

$$x = \Re\left(\sum_{l=1}^{L} Z_K(l\omega)\bar{c}_l\right), \quad y = x + w \tag{18}$$

with $w \sim \mathcal{N}(0, \sigma^2 I)$. A real-valued $y \in \mathbb{R}^N$ multi-pitch SDP estimator can be obtained from the complex-valued SDP estimator (17) as

$$\begin{array}{ll} \underset{X_{11},X_{12},X_{22}}{\text{minimize}} & \frac{1}{2}(\mathbf{tr}(X_{11}) + \mathbf{tr}(X_{22})) \\ \text{subject to} & \|y - P\,\mathbf{vec}(\Re(X_{12}))\|_2 \leq \delta \\ & \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^H & X_{22} \end{bmatrix} \succeq 0 \\ & X_{11} \in \mathbb{T}^{NL}, X_{12} \in \mathbb{C}^{NL \times L}, X_{22} \in \mathbb{H}^L \end{array} \tag{19}$$

with a solution $(X_{11}^\star, X_{22}^\star, X_{12}^\star)$. The problem (19) can be formulated as a real SDP via the following observations: the optimal objective is $\frac{1}{2}(\mathbf{tr}(X_{11}^\star) + \mathbf{tr}(X_{22}^\star)) = \frac{1}{2}(\mathbf{tr}(\Re(X_{11}^\star)) + \mathbf{tr}(\Re(X_{22}^\star))$ and

$$\begin{bmatrix} X_{11}^\star & X_{12}^\star \\ (X_{12}^\star)^H & X_{22}^\star \end{bmatrix} \succeq 0 \Rightarrow \Re\left(\begin{bmatrix} X_{11}^\star & X_{12}^\star \\ (X_{12}^\star)^H & X_{22}^\star \end{bmatrix}\right) \succeq 0\,. \tag{20}$$

Further, if $X_{11}^\star$ is Toeplitz, then $\Re(X_{11}^\star)$ is also Toeplitz. So $(\Re(X_{11}^\star), \Re(X_{22}^\star), \Re(X_{12}^\star))$ is also a solution to (19). Then we may instead of the complex SDP (19) formulate the equivalent real SDP

$$\begin{array}{ll} \underset{X_{11},X_{12},X_{22}}{\text{minimize}} & \frac{1}{2}(\mathbf{tr}(X_{11}) + \mathbf{tr}(X_{22})) \\ \text{subject to} & \|y - P\,\mathbf{vec}(X_{12})\|_2 \leq \delta \\ & \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \end{bmatrix} \succeq 0 \\ & X_{11} \in \mathbb{S}^{NL} \cap \mathbb{T}^{NL}, X_{12} \in \mathbb{R}^{NL \times L}, X_{22} \in \mathbb{S}^L \end{array} \tag{21}$$

with a solution that is also a solution of (19). Notice that the real SDP is of the same dimension but half the number of (real) variables. We will denote an estimator based on (21) a SDP multi-pitch (SDPMP) estimator. Based on the Carathéodory parameterization (see e.g. [15]), the matrix $X_{11} \in \mathbb{T}^{NL} \cap \mathbb{S}^{NL}$ will contain pairs of symmetric frequencies $\omega_k = -\omega_{k+1}$ and possible a single $\omega_k = 0$.

5. FREQUENCY CONSTRAINT

If the signal y is Nyquist sampled then the fundamentals obeys $-\pi \leq L\omega_k \leq \pi$. The frequencies in the dictionary $\mathbb{A}_n$ can be constrained using the parameters α, β by adding the following semidefinite cone constraint [10] to e.g. (21):

$$-e^{j\alpha} F X_{11} G^T - e^{-j\alpha} G X_{11} F^T + 2\cos(\beta) G X_{11} G^T \preceq 0 \tag{22}$$

where $F = \begin{bmatrix} 0 & I_{NL-1} \end{bmatrix}, G = \begin{bmatrix} I_{NL-1} & 0 \end{bmatrix}$. Then with the selection $\alpha = 0$, (22) is a real semidefinite cone constraint

$$F X_{11} G^T + G X_{11} F^T - 2\cos(\beta) G X_{11} G^T \succeq 0\,. \tag{23}$$

A symmetric Toeplitz matrix can be formed using the function

$$T(z) = \begin{bmatrix} z_0 & z_1 & \cdots & z_n \\ z_1 & z_0 & \cdots & z_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ z_n & z_{n-1} & \cdots & z_0 \end{bmatrix}. \tag{24}$$

If $X_{11} \in \mathbb{S}^{NL} \cap \mathbb{T}^{NL}$ then let x_{11} be the first column of X_{11} such that $T(x_{11}) = X_{11}$, and the left-hand side of (23) can be written as

$$T\left(\begin{bmatrix} 2t_1 \\ t_0 + t_2 \\ \vdots \\ t_{NL-3} + t_{NL-1} \end{bmatrix}\right) - 2\cos(\beta) T\left(\begin{bmatrix} t_0 \\ t_1 \\ \vdots \\ t_{NL-2} \end{bmatrix}\right) = T(C_\beta x_{11}) \tag{25}$$

with $t = x_{11}$ and an implicit definition of the matrix C_β. That is, the left-hand side of (23) is also Toeplitz. The frequency constrained SDPMP (CSDPMP) estimator with $\beta = \pi/L$ is

$$\begin{array}{ll} \underset{K,\omega_k,c_k}{\text{minimize}} & \sum_{k=1}^{K} \|c_k\|_2 \\ \text{subject to} & \|y - x\|_2 \leq \delta \\ & x = \Re\left(\sum_{l=1}^{L}\sum_{k=1}^{K} z_N(l\omega_k)[c_k]_l\right) \\ & |\omega_k| \leq \frac{\pi}{L}, k = 1, \ldots, K \end{array} \tag{26}$$

which can be formulated as the SDP

$$\begin{array}{ll} \underset{X_{11},X_{12},X_{22}}{\text{minimize}} & \frac{1}{2}(\mathbf{tr}(T(x_{11})) + \mathbf{tr}(X_{22})) \\ \text{subject to} & \|y - P\,\mathbf{vec}(X_{12})\|_2 \leq \delta \\ & \begin{bmatrix} T(x_{11}) & X_{12} \\ X_{12}^T & X_{22} \end{bmatrix} \succeq 0 \\ & T(C_{\pi/L} x_{11}) \succeq 0 \\ & x_{11} \in \mathbb{R}^{NL}, X_{22} \in \mathbb{S}^L, X_{12} \in \mathbb{R}^{NL \times L}\,. \end{array} \tag{27}$$

The matrix $X_{11} = T(x_{11}) \in \mathbb{S}^{NL}$ will again contain pairs of symmetric frequencies $\omega_k = -\omega_{k+1}$ but with $|\omega_k| = |\omega_{k+1}| \leq \beta = \pi/L$ and possibly a single $\omega_k = 0$.

6. SIMULATIONS

We will in the following simulations investigate resilience to spectral smearing caused by closely spaced harmonics and noise with known model order K. We perform Monte Carlo simulations using $R = 500$ repetitions of the real-valued AWGN channel (18). The proposed SDPMP and CSDPMP estimators are implemented with a custom solver for CVXOPT [16] to exploit the Toeplitz structure as presented in [17][1]. We use the "noiseless" ESPRIT algorithm outlined in [11] to extract the frequencies from $X_{11}^\star$ (compare with [15, pp. 174–175]). The final estimates are obtained in two steps: *1)* solve the SDP with the regularization parameter δ selected by averaging the smallest $\frac{1}{3}$ of the coefficients of the periodogram *2)* extract the frequencies $\omega^\star$, re-select the regularization parameter as $\delta = \min_{\bar{c}_l} \|y - \Re\left(\sum_{l=1}^{L} Z_K(l\omega^\star)\bar{c}_l\right)\|_2$ using linear least-squares and re-solve the SDP. We will compare with methods and implementations from [1]: *i)* an approximate NLS method (ANLS) based on the harmonic summation method *ii)* a subspace method (ORTH) based on MUSIC *iii)* an optimal filtering method (OPTFILT) based on the

[1] Estimators and simulation code available at `kom.aau.dk/~tlj`

Capon approach. These require complex-valued data and we form these using the Hilbert transform and obtain $\bar{y} \in \mathbb{C}^{N/2}$. This mapping does not introduce a significant error if the frequencies are not too close to 0 and π [1]. In our case with $N = 160$, the lowest frequency is $0.1580 \approx 2\pi(4/N)$ and the highest is $0.6364 \approx 2\pi(16/N)$. We will not compare with the method in [5] since a key element of this method is incorporation of model order selection and does not lend itself easily to the case of known model orders.

If two pitches are well separated and not too low such that there is no significant overlap of the harmonics (see e.g. the discussion in [1]), then the accuracy of estimating the pitches should at-least for unbiased estimators be governed by the asymptotic CRLB for estimating a single pitch $\hat{\omega}_k$:

$$\mathbf{var}(\hat{\omega}_k) \geq \frac{24\sigma^2}{(N(N^2-1))\sum_{l=1}^{L} A_{k,l}^2 l^2} \tag{28}$$

where $A_{k,l} = |[\bar{c}_l]_k|$ is the amplitude. Notice that this bound depends on the "enhanced SNR" [18] (for a single pitch) or pseudo SNR (PSNR) for the kth pitch [4]

$$\mathrm{PSNR}_k = 10\log_{10}\frac{\sum_{l=1}^{L} A_{k,l}^2 l^2}{\sigma^2}\,. \tag{29}$$

In the following two experiments we will assess the accuracy of the methods based on the root-mean-squared-error

$$\mathrm{RMSE} = \sqrt{\frac{1}{RK}\sum_{r=1}^{R}\sum_{k=1}^{K}|\omega_{k,r} - \hat{\omega}_{k,r}|^2} \tag{30}$$

where $\omega_{k,r}$ denotes the kth fundamental for the rth realization and $\hat{\omega}_{k,r}$ denotes an estimate. In the first experiment we will consider the difference between the fundamental frequencies of $K = 2$ sources

$$\omega_2 - \omega_1 = \Delta\,. \tag{31}$$

We select the first fundamental from a uniform distribution $\omega_1 \sim 2\pi\mathcal{U}(7/N, 14/N)$ and for each realization $\omega_2 = \omega_1 + \Delta$. We let $N = 160$, $L = 3$ and $A_{k,l}^2 = 1$ with *i.i.d.* uniform phase $\mathcal{U}(0, 2\pi)$ as in [4]. The result is shown in Fig. 1. We observe that as Δ increases the SDPMP, CSDPMP, ORTH and OPTFILT approaches the CRLB as expected and for the latter two as reported in [4]. Further, we observe that CSDPMP gives state-of-the-art performance for separating two closely spaced pitches. This indicates that CSDPMP may offer a higher time-frequency resolution [2].

In the second experiment we will investigate the accuracy as a function of the PSNR with two pitches fixed at $\omega_1 = 0.1580$ and $\omega_2 = 0.6364$ as in [4] selected due to the near integer relationship. We observe in Fig. 2 that using the frequency constrained CSDPMP gives about the same RMSE accuracy as the SDPMP and that both methods almost achieve the CRLB. The other methods show a slightly larger gap to the CRLB and/or approaches the bound at a larger PSNR.

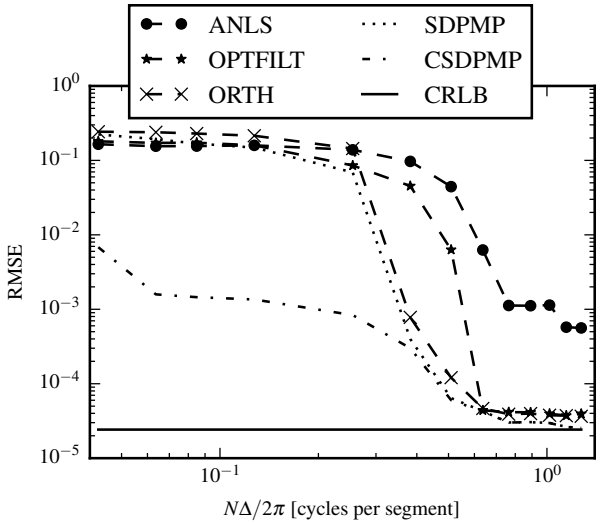

Fig. 1. RMSE as a function of the fundamental frequency difference $\omega_2 - \omega_1 = \Delta$, $K = 2$, $N = 160$, $L = 3$, $\mathrm{PSNR}_1 = \mathrm{PSNR}_2 = 40$ [dB].

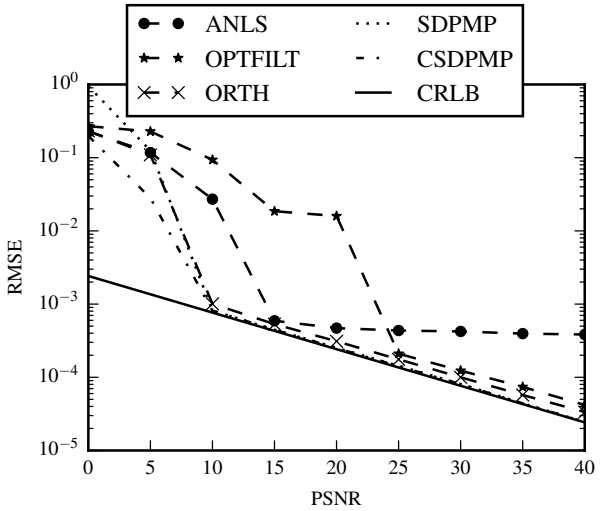

Fig. 2. RMSE as a function of the $\mathrm{PSNR} = \mathrm{PSNR}_1 = \mathrm{PSNR}_2$, $K = 2$, $N = 160$, $L = 3$, and $\omega_1 = 0.1580$, $\omega_2 = 0.6364$.

7. CONCLUSION

In this paper we have formulated a real-valued multi-pitch estimator based on continuous sparse optimization. This type of framework has been studied in superresolution, gridless compressed sensing, and atomic norm optimization where complex-valued data models have been investigated. Many applications will naturally have real-valued data and we show how the standard complex-valued data formulation can be converted to real-valued data using a real semidefinite program with the same dimensions. Further, if the signal is Nyquist sampled, then we should impose a frequency constraint on the fundamentals in order to achieve state-of-the-art estimation accuracy.

8. REFERENCES

[1] M. G. Christensen and A. Jakobsson, *Multi-Pitch Estimation*, San Rafael, CA, USA: Morgan & Claypool, 2009.

[2] D. Chazan, Y. Stettiner, and D. Malah, "Optimal multi-pitch estimation using the EM algorithm for co-channel speech separation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, April 1993, vol. 2, pp. 728–731.

[3] P. Stoica, G. Tangy, Z. Yangz, and D. Zachariah, "Gridless compressive-sensing methods for frequency estimation: Points of tangency and links to basics," in *Proc. European Signal Processing Conf.*, Sept 2014, pp. 1831–1835.

[4] M. G. Christensen, P. Stoica, A. Jakobsson, and S. H. Jensen, "Multi-pitch estimation," *Signal Processing*, vol. 88, no. 4, pp. 972–983, Apr. 2008.

[5] F. Elvander, T. Kronvall, S. I. Adalbjörnsson, and A. Jakobsson, "An adaptive penalty multi-pitch estimator with self-regularization," *Signal Processing*, vol. 127, pp. 56 – 70, 2016.

[6] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Found. Comp. Math.*, vol. 12, no. 6, pp. 805–849, 2012.

[7] E. J. Candès and C. Fernandez-Granda, "Super-resolution from noisy data," *J. Fourier Anal. Appl.*, vol. 19, no. 6, pp. 1229–1254, 2013.

[8] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, Nov 2013.

[9] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5987–5999, Dec 2013.

[10] H.-H. Chao and L. Vandenberghe, "Extension of semidefinite programming methods for atomic decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2016, pp. 4757–4761.

[11] H.-H. Chao and L. Vandenberghe, "Semidefinite representations of gauge functions for structured low-rank matrix decomposition," *ArXiv e-prints*, Apr. 2016.

[12] Z. Yang, L. Xie, and C. Zhang, "A discretization-free sparse and parametric approach for linear array signal processing," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 4959–4973, Oct 2014.

[13] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1257–1269, March 2016.

[14] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3139–3153, June 2015.

[15] P. Stoica and R. L. Moses, *Spectral analysis of signals*, Pearson/Prentice Hall, 2005.

[16] M. S. Andersen, J. Dahl, Z. Liu, and L. Vandenberghe, "Interior-point methods for large-scale cone programming," in *Optimization for Machine Learning*, S. Sra, S. Nowozin, and S. J. Wright, Eds. MIT Press, 2011.

[17] T. Roh and L. Vandenberghe, "Discrete transforms, semidefinite programming and sum-of-squares representations of nonnegative polynomials," *SIAM J. Optimiz.*, vol. 16, pp. 939–964, 2006.

[18] A. Nehorai and B. Porat, "Adaptive comb filtering for harmonic signal enhancement," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 5, pp. 1124–1138, Oct. 1986.